



LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-TR-643914

Assessment of Land Surface Model Performance in WRF for Simulating Wind at Heights Relevant to the Wind Energy Community

S. Wharton, M. Simpson, J. Osuna, J. Newman,
S. Biraud

September 17, 2013

Disclaimer

This document was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor Lawrence Livermore National Security, LLC, nor any of their employees makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or Lawrence Livermore National Security, LLC. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or Lawrence Livermore National Security, LLC, and shall not be used for advertising or product endorsement purposes.

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

Assessment of Land Surface Model Performance in WRF for Simulating Wind at Heights Relevant to the Wind Energy Community

Sonia Wharton¹, Matthew Simpson¹, Jessica Osuna¹, Jennifer Newman², Sebastian Biraud³

¹Atmospheric, Earth and Energy Division, Lawrence Livermore National Laboratory, Livermore, CA

²School of Meteorology, University of Oklahoma, Norman, OK

³Earth Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA

Abstract

The Weather Research and Forecasting (WRF) model was used to investigate the impact of land surface model (LSM) physics on the simulated surface energy balance and the near surface wind profile, including heights reached by a modern wind turbine. Simulations were made for a domain centered over the Department of Energy (DOE) Southern Great Plains (SGP) Atmospheric Radiation Measurement Program (ARM) Central Facility in northern Oklahoma. High-resolution measurements of surface fluxes and profiles of wind speed, direction and turbulence were available for model validation. We tested 10 common LSM configurations in WRF including Noah, Noah-MP, SSiB, Pleim-Xiu, RUC, and others to evaluate (1) the accuracy of simulated surface energy fluxes, (2) the accuracy of forecasted rotor-disk height wind speeds, and (3) the sensitivity of rotor disk wind speed to the choice of LSM. WRF was run for four, two-week periods covering different canopy, surface meteorology and soil moisture conditions. LSM performances varied by simulation period and errors were greatest during periods with high soil moisture and live crop cover, indicating highest LSM sensitivity when PBL behavior is most coupled to land-atmosphere exchange. Higher wind speed errors also occurred at night during the spring, possibly from the difficulty with simulating nocturnal low-level jets. We found that choice of land surface model led to a ~10% improvement in simulating hub-height wind speed. Errors in wind shear were also sensitive to the choice of LSM. Overall the best performing models were Noah and Noah-MP. The variability of LSM performance across different soil moisture and vegetation canopy conditions suggests LSM representation of surface energy exchange processes in WRF remains a large source of model uncertainty. Furthermore, our results indicate the choice of LSM impacts simulated wind fields in WRF at heights relevant to the wind industry, at least for the Southern Great Plains region.

1. Introduction

Wind power forecasting continues to have difficulties in accurately predicting the occurrence and intensity of atmospheric conditions at the heights spanned by industrial-scale wind turbines (~ 40 to 200 m above ground level, AGL). Identifying areas in numerical models which may lead to a wind forecasting improvement is a warranted endeavor as a forecasting improvement of as little as 10-20% could result in hundreds of millions of dollars (\$140-260M) in annual operating cost savings for the U.S. wind industry (Lew et al. 2011). One such candidate is the user's choice of land surface model (LSM) used in numerical weather prediction models. The Weather Research and Forecasting (WRF) model (Shamarock et al. 2008) is heavily utilized in the atmospheric community for weather prediction and more recently for wind forecasting, including its use in predicting hub-height wind speed and direction in the wind power industry. WRF has many physics options that are user defined, including options for choosing the planetary boundary layer (PBL) scheme, surface layer scheme, radiation scheme, microphysics and convective schemes, and land surface model. It is the choice of LSM that we examine in detail for assessing it as a candidate area for wind forecasting improvement. Given the large number of currently available LSMs in WRF, it remains a challenge for modelers to select the appropriate land surface scheme to fit their needs (Chen et al. 2012). We hypothesize that wind speed at heights equivalent to a turbine rotor disk are sensitive to LSM choice due to variations in the sophistication of LSM calculation of canopy layer energy fluxes. This theory is tested for the Southern Great Plains region of the U.S.

In recent years the Department of Energy (DOE) Southern Great Plains (SGP) Atmospheric Radiation Measurement Program (ARM) Central Facility in northern Oklahoma has become an ideal location for conducting wind energy research as multi-MW wind farms are currently being installed in great numbers around the well instrumented DOE site. In 2012, a multi-MW wind farm called Chisholm View began operation to the southwest of the ARM Central Facility. The nearest turbines to the SGP ARM Central Facility are approximately a couple of miles away. Chisholm View, at present, has 140 operating GE 1.6 MW turbines, each with a hub-height of 80 m, and rotor diameter of 82.5 m. Each turbine blade tip has a minimum height of 39 m and a maximum height of 121 m.

Energy exchange at the surface largely drives lower PBL flow features in the daytime, as thermally-generated, buoyant sensible heat fluxes, in combination with mesoscale forcing and surface roughness, largely determine the magnitudes of wind shear and turbulence at heights encountered by a wind turbine rotor-disk (Wharton and Lundquist 2012b). While studies have examined the role of surface and subsurface properties on near-surface meteorology and PBL development (e.g., Anderson and Verma 1986, Baldocchi 1994, Betts et al. 1996, Raupach 2000, Patton et al. 2004, Desai et al. 2006, Maxwell et al. 2007, Fischer et al. 2007.), as well as the role of the PBL scheme in WRF for wind forecasting (e.g., Storm and Basu 2010, Draxl et al. 2012), few studies have specifically looked at the role of surface energy exchange and LSM choice on rotor-disk wind profiles and LSM importance in WRF for wind forecasting. This may be a potential area for wind forecasting improvement as numerous studies have shown that wind shear and turbulence characteristics are dependent on stability and have significant impacts on turbine power generation (e.g., van den Berg 2008, Wagner et al. 2009, Wharton and Lundquist. 2012a,b).

Despite the significant influence of land-atmosphere exchange on the daytime PBL, uncertainty remains in the parameterization of surface heat and moisture fluxes in numerical weather models (Trier et al. 2010). The LSM in WRF provides heat and moisture fluxes over land to provide a lower boundary condition for vertical transport in the PBL scheme. These fluxes of sensible heat and latent energy are dependent on the surface meteorology, radiative forcing, soil properties and land use type. Thus, an accurate description of the land surface and vegetation characteristics is needed in any numerical weather prediction model. WRF is an ideal model for studying the impact of the simulated energy balance on near surface winds for several reasons. Sub-grid scale processes are represented in WRF by physics schemes that parameterize atmospheric phenomena such as turbulent mixing, surface energy transfer, and cloud microphysics. Due to the availability of numerous physics model choices, users are able to customize the WRF physics configuration to match their needs. Numerous LSM options are available in WRF that range from simplistic treatment of sub-surface processes to sophisticated physical models.

Surface energy exchange is determined by the terrestrial radiation budget, defined in Eq. 1, where R_n is the net radiation flux ($W m^{-2}$), H is the sensible heat flux ($W m^{-2}$), LE is the latent energy flux ($W m^{-2}$), G is the ground heat flux ($W m^{-2}$) and S is the biomass storage heat flux ($W m^{-2}$). While some energy is absorbed by the ground and biomass ($G + S$), this is on average less than 15% for most plant canopies, and the majority of available energy will be transferred back into the atmosphere as sensible and latent heat (Betts et al. 1996). Latent heat is the quantity of heat absorbed or released by water undergoing a change of state. Over a plant canopy, LE is most often the heat released by water as it changes from a liquid to gaseous state through evapotranspiration (evaporative water loss from the vegetation and soil plus water loss through plant transpiration).

$$R_n - H - LE - G - S = 0 \quad \text{Eq. 1}$$

The exact magnitudes of H , LE , G and S are dependent on many variables, including the amount of incoming radiation, soil moisture availability, groundwater availability and plant access, soil properties, canopy properties (e.g., species, albedo, biomass density, leaf area index, rooting depth), temperature, and entrainment of dry air into the boundary layer from the free atmosphere. The ratio of sensible heat to latent energy transfer is called the Bowen ratio and is usually expressed as a midday average. A higher Bowen ratio indicates that a larger portion of available energy is partitioned into sensible heat than latent heat. Over a desert, for example, one would expect a high Bowen ratio (>10), while over a tropical forest it may be less than 1 due to high evapotranspiration.

Studies such as CASES-97 and IHOP_2002 have shown that latent and sensible heat fluxes in the Southern Great Plains of the U.S. vary significantly in time and space due to land use, climatology and soil type differences across the landscape (LeMone et al. 2000, Weckwerth et al. 2004, LeMone et al. 2006). Standard WRF does not capture fine scale land use heterogeneity nor does it usually capture fine-scale temporal vegetation changes. Furthermore, LSMs in WRF use a seasonally-adjusted or annually-fixed land use value from look-up tables to assign surface characteristics, but crop cover can change frequently and suddenly in this region; for example, rapid canopy changes would be caused by crop

“greenup” or senescence, flood or hail crop losses, or harvest or field tillage events. This coarse approach for incorporating land use and vegetation property changes in LSMs is likely a shortcoming in areas such as the Great Plains with high spatial and temporal surface variability. To improve modeling in heterogeneous areas, some groups have incorporated high-resolution remote sensing surface data into their land surface model (e.g. Miller et al. 2006, Case et al. 2012). This “mosaic approach” using satellite data assimilation can be incorporated into certain WRF-coupled LSMs such as Noah, RUC and the Community Land Model (CLM).

While satellite data assimilation of near real-time surface properties would likely result in more accurate surface fluxes, and potentially more accurate simulations of turbine-relevant wind fields, this was not done here for the simple reason that we wanted to evaluate the accuracy of each LSM in WRF by running WRF in a “standard” or “operational” mode as a wind plant operator or forecaster would do. In this manner, we hope our conclusions will be more beneficial to the wind industry and energy utilities communities. The recent operation of the neighboring Chisholm wind farm provided motivation for studying the importance of LSMs in WRF for accurate simulations of wind speed and shear at heights encountered by modern wind turbines. Energy flux and surface meteorology observations from nearby SGP ARM were used for identifying the physics that govern lower planetary boundary layer behavior and for validating the ten WRF ensemble members. Specific project goals were to quantify: (1) the accuracy of simulated surface energy fluxes to flux tower measurements, (2) the accuracy of simulated wind speeds to observations at rotor-disk heights, and (3) the sensitivity of simulated hub-height rotor disk wind speed to the choice of land surface model in the Southern Great Plains region. This study is the first to our knowledge to systematically evaluate the role of LSMs in WRF for wind industry applications.

2. Methods

2a. Site Description and Instrumentation

All surface and atmospheric measurements were taken at the DOE SGP ARM Central Facility near Lamont, Oklahoma (36.61°N, 97.49°W). SGP ARM is a state-of-the-art research facility for atmospheric, climate, and radiation measurements and is set in a rural landscape dominated by pastureland and annual crops (Figure 1). The landscape is relatively flat and local terrain slope is less than three degrees.

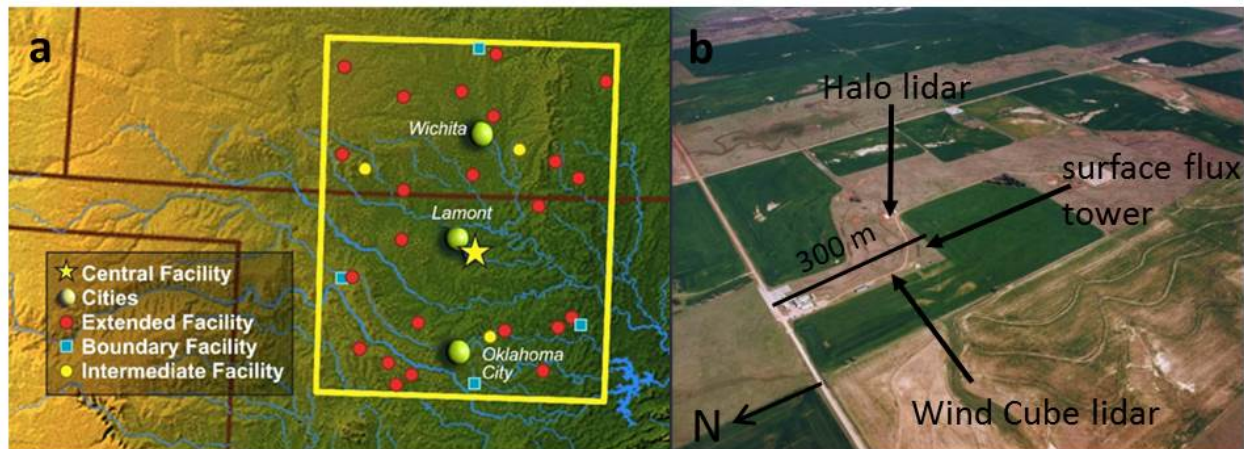


Figure 1 (a) Map showing the boundaries of the greater Southern Great Plains (SGP) Atmospheric Radiation Measurement (ARM) Facility in Oklahoma and Kansas (yellow box) including the Central Facility in Northern Oklahoma. **(b)** Aerial photograph over the Central Facility depicting the fine scale spatial heterogeneity in the area and locations of instruments used in this study. Live annual crops are indicated by dark green while pasturelands and fallow fields are light brown.

The species of annual crops planted in fields adjacent to the Central Facility varies from year to year and season to season depending on climatic conditions and market prices and can include winter wheat and canola, typically grown from fall to early summer, and the summer crops corn, sorghum, cowpeas, barley and soybeans. The region has fine scale spatial heterogeneity due to small land holdings and changing crop rotations (Torn et al. 2011), as indicated in Figure 1b. Regional heterogeneity is further enhanced by climatic gradients which change precipitation amounts across the larger SGP ARM domain from the northwest (380 mm) to the southeast (1270 mm) (Figure 1a). The crops are not irrigated in the area and crop losses, delayed plantings, and early or late harvests can result from anomalous climate conditions. Area soils are classified as well-drained silt loam, silty clay loam or clay loam. The permanent wilting point for these soil types are $0.10 \text{ m}^3 \text{ m}^{-3}$, $0.13 \text{ m}^3 \text{ m}^{-3}$, $0.15 \text{ m}^3 \text{ m}^{-3}$, respectively. Permanent wilting point is defined as the water content of the soil when most plants will wilt and fail to recover their turgor even if the soil moisture is replenished.

In our study we utilized measurements from a 4 m tall surface flux station as described in Billesbach et al. 2004, Fischer et al. 2007 and Riley et al. 2009. The surface flux tower faces south, the predominant wind direction except during winter, and measures mass and energy fluxes over an annual crop field. The surface flux tower has an open-path infrared gas analyzer (LICOR Li-7500) that measures

atmospheric carbon dioxide and water vapor (i.e., latent energy flux) and a 3D sonic anemometer (Gill-Solent WindMaster Pro, Gill) that measures the wind vectors and sonic temperature, from which the sensible heat flux is calculated. Other standard meteorology measurements including air temperature, relative humidity, the four-stream energy flux (incoming and outgoing shortwave and longwave radiation), soil heat flux, soil moisture (average of upper and lower profiles), soil temperature and precipitation are also measured. The energy fluxes were calculated using a 30-minute averaging period and were post-processed using standard algorithms for spike removal, density corrections (WPL corrections, Webb et al. 1980), spectral corrections (Moore 1986), coordinate rotation and more. The flux tower footprint is on the order of a couple of hundred meters. Further flux processing details can be found in Riley et al. 2009. Crop species present during the study are shown later in Figure 5. The flux tower and surface meteorology measurements were downloaded from the ARM archive: <http://www.archive.arm.gov/>.

Profiles of wind speed and direction were measured using a Light Detection and Ranging (lidar) instrument. The SGP ARM Central Facility has a continuously running Halo scanning lidar (Halo Photonics, Worcestershire, UK) which provided measurements of horizontal wind speed and direction, taken once an hour in 2011, for heights ranging from 75 m up to up to 10 km (Newsom 2012). The Halo lidar provides wind measurements roughly every 10 m in the first 200 m above the surface. In November 2012, we installed a vertically-facing Wind Cube v2 lidar system (Leosphere, Orsay, France) which measured the wind speed and direction roughly every 5-15 m from 50 m to 180 m above the surface at a 1 Hz rate. The Wind Cube v2 lidar has the ability to provide higher spatial-resolution close to the surface. The Wind Cube measurements were averaged over ten-minute periods. Using the Halo and Wind Cube lidars we calculated wind speed at a height equivalent to a turbine nacelle (referred to as the hub-height wind speed; 80 m AGL) as well as wind shear across heights equivalent to a nearby turbine rotor disk (39 m to 121 m AGL).

2b. Site Climatology

The climate is continental to semi-arid and the area is prone to moderate and extreme drought (Torn et al. 2011). Local air temperature and precipitation data were downloaded from the Oklahoma Mesonet (<http://www.mesonet.org/>) for four stations nearest to the SGP ARM Central Facility in order to include local weather variability, which can be high due to convective storms. Long term (2002-2012) mean air temperature (T_a) was 15.0 °C and mean precipitation (PPT) was 812 mm. In comparison, the WRF simulation years were warmer and drier than the 10-year norm. Mean annual air temperature was 15.3 °C in 2011 and 16.7 °C in 2012; annual precipitation was 641 mm in 2011 and 575 mm in 2012. Drier than normal conditions occurred in spring of 2011 and a prolonged drought occurred in summer and autumn of 2012. Winds at the site were generally from the south, particularly in the summer and autumn, with a dominant northerly component also present in the winter and spring months. Annual mean wind speed was just over 7.5 m/s at a height of 60 m at the tall meteorological tower from 2010-2012. Wind speed varied from month to month with the windiest months occurring during spring (March-June) and autumn (October-November), while the summer months (July-September) had on average lower wind speeds (Figure 2).

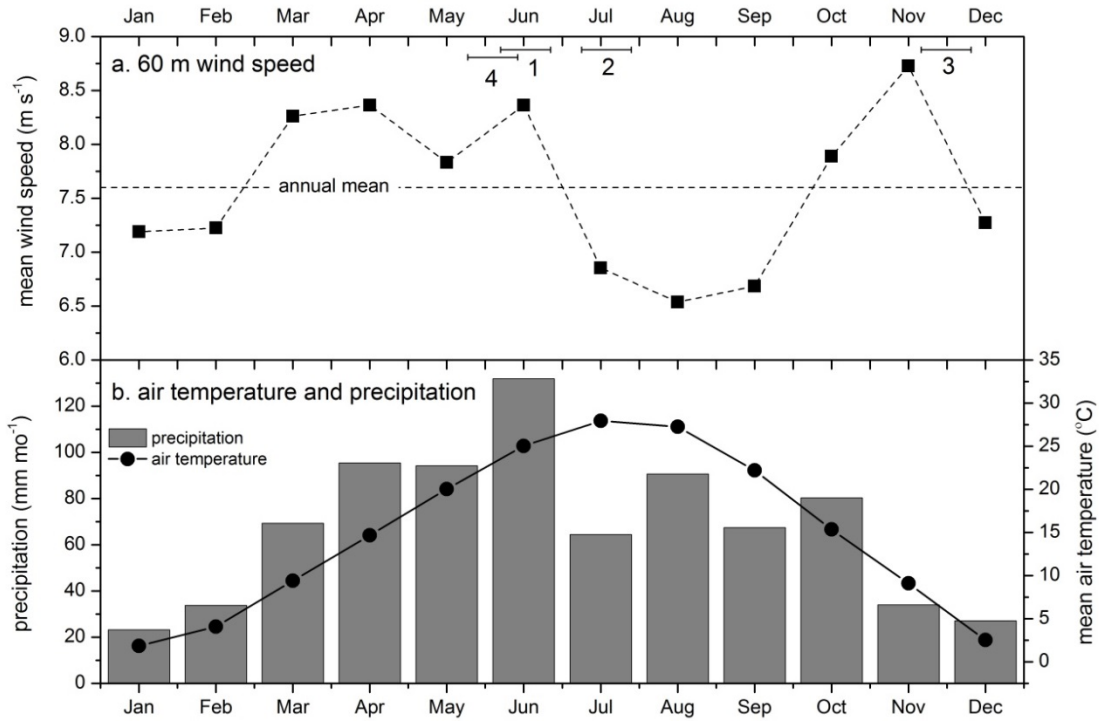


Figure 2 (a) Mean monthly wind speed at 60 m from 2010-2012 at the SGP ARM Central Facility. (b) Long-term mean monthly air temperature and total precipitation from 2002-2012. Also plotted is time of year for the WRF simulation periods 1-4 as described in Section 2c.

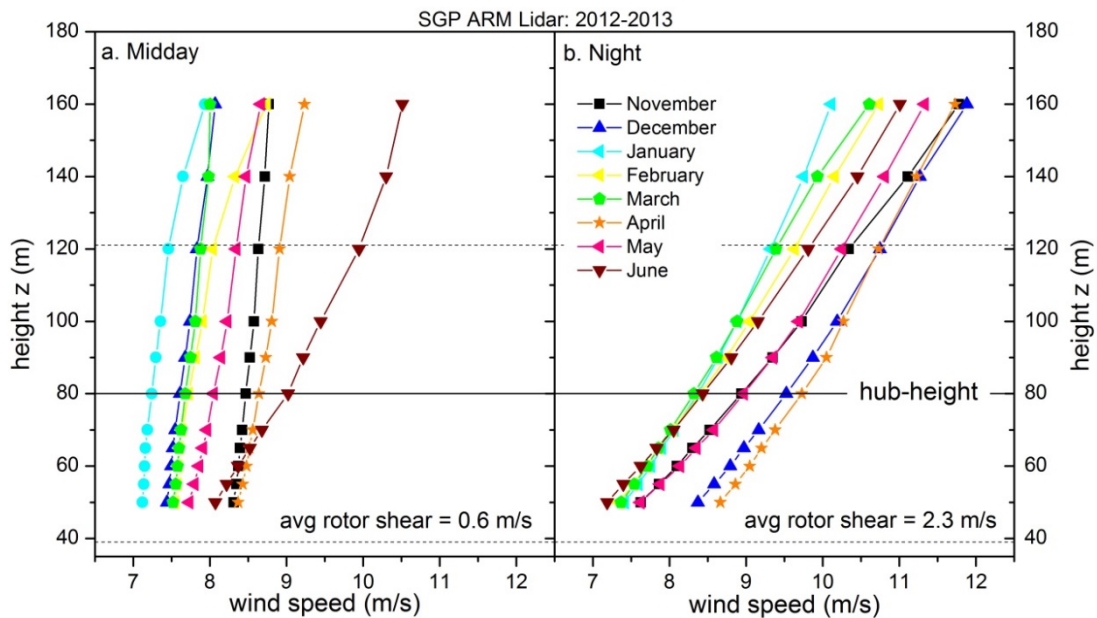


Figure 3 Mean wind speed profiles by month for (a) midday and (b) nighttime hours during the Wind Cube v2 field campaign (November 2012 – June 2013). These profiles show average wind shear conditions across heights which an 80 m hub-height wind turbine would encounter (dotted lines indicate heights equivalent to a rotor disk).

Seasonal-variability in surface-layer wind shear was greater during midday hours than nighttime hours as measured by the lidar (Figure 3). Nighttime wind shear was overall higher than daytime and was on average 2.3 m/s between 120 m and 50 m versus 0.6 m/s during midday. Midday wind shear showed significantly greater wind shear during June 2013 than any other month during the field campaign. This suggests that June may have a very different energy partitioning ratio than the other months.

2c. Simulation periods

To study the impact of LSM model choice on WRF simulated near surface wind profiles, we performed atmospheric simulations over a variety of land surface conditions that would affect the surface energy budget, including times of moist soil, dry soil, live crop canopy and senesced crop canopy (Table 1):

Table 1 Brief description of simulation periods highlighting the canopy and meteorology differences.

	Simulation 1	Simulation 2	Simulation 3	Simulation 4
Dates	June 10 - 23, 2011	July 13 - 27, 2011	Nov 23 - Dec 7, 2012	May 24 - June 10, 2013
Characteristics	very warm, wet	hot, dry	cool, dry	warm, wet
Mean air temperature	27.2 °C	32.1 °C	10.2 °C	20.7 °C
PPT, including the prior two weeks	55.8 mm	8.8 mm	18.1 mm	138.6 mm
Mean soil moisture	24%	6%	9%	22%
Crop type	canola	None	planted wheat	Wheat
Field cover	peak LAI, live canopy	bare soil + crop residual	crop residual + few live seedlings	peak LAI, live canopy

Actual conditions in the crop field during June 2011 included a live canopy with peak Leaf Area Index (LAI) in early June followed by senescence during the next two weeks up until the crop was harvested (Figure 4a). The field was tilled on June 30, and actual conditions during the July period included a bare, brown field with some dead crop residual (Figure 4b). At this time the field and the surrounding area were covered with <5% photosynthesizing vegetation and the soils were bare and very dry. In fall 2012, the local area was a mixture of <0.5 m senesced grasses (in the pasturelands) and dead cowpea crop residual (< 1 cm) and bare ground in the field with a few emerging wheat seedlings (Figure 4c). November also included cool temperatures and very dry soils and was considered the “control” period as energy fluxes would be relatively low during this time of year and not have a significant influence on PBL behavior and development. The four and final period, May-June 2013, included a green, peak LAI wheat canopy with high soil moisture and high evapotranspiration which rapidly senesced in mid-June (Figure 4d).

Figure 5 shows a time series of shortwave albedo (clear sky fitted), used as a proxy for “canopy greenness” as well as the exact timing of planting, harvest and tilling events. Lower albedo measurements are interpreted as a “greener” canopy where the likelihood of large latent energy fluxes is high because the vegetation is actively taking up water and transpiring. Higher albedo measurements are interpreted as a senesced canopy or a combination of dead vegetation plus bare soil. High albedo results in a smaller amount of energy available for both latent and sensible heat as a greater portion of

incoming shortwave radiation is reflected back into the atmosphere. If high albedo is a result of a dead canopy or bare, dry soil, this further suggests that the latent energy fluxes will be very low.

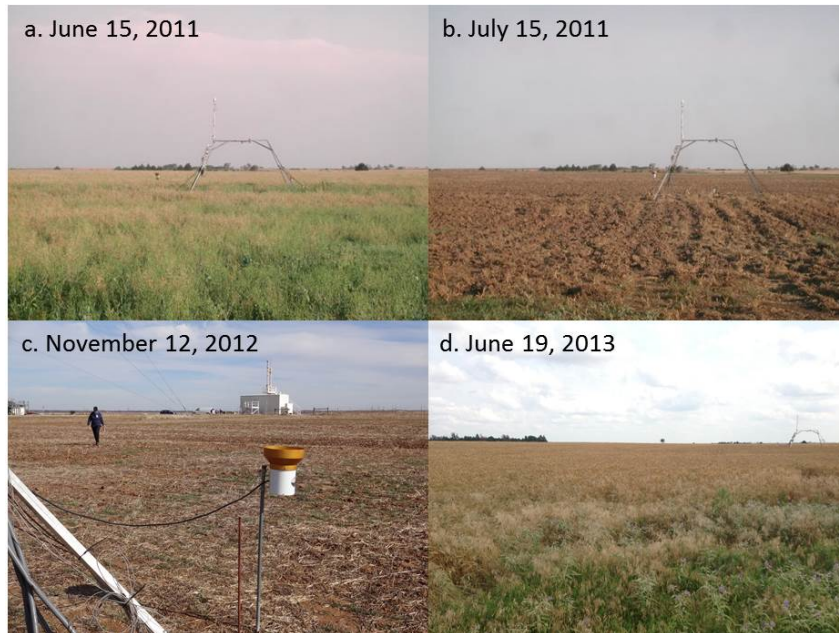


Figure 4 Photographs showing the canopy cover in the crop field at the Southern Great Plains ARM site on dates during or near to the four WRF simulation periods.

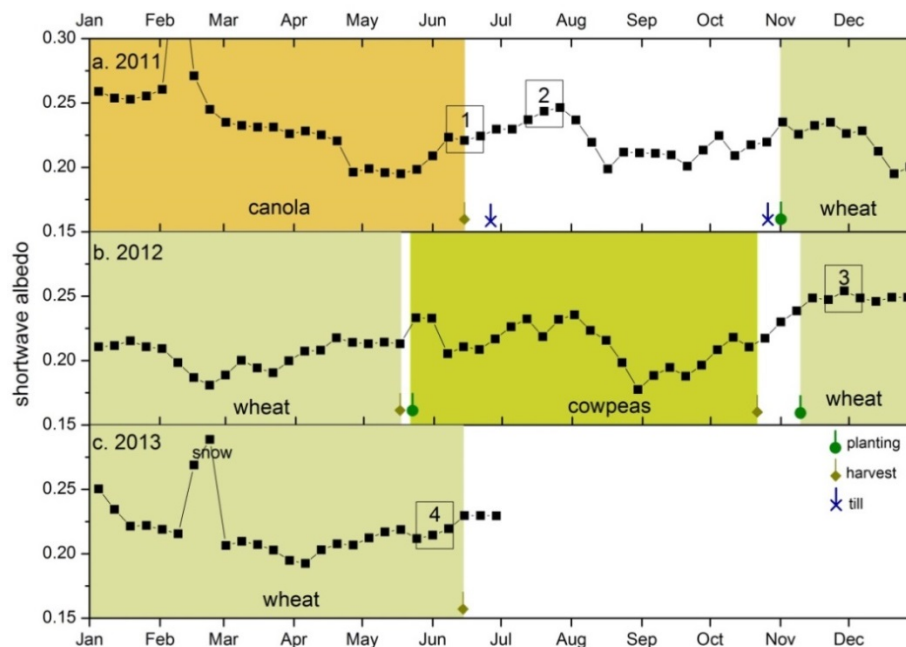


Figure 5 Time series of mean weekly shortwave albedo at the SGP ARM Central Facility by year and the planting, harvest, and till schedule in the flux tower field. The numbers indicate the simulation periods. Higher albedo could indicate either bare ground or a senesced canopy. High albedo spikes are likely due to snow cover. Lower albedo likely indicates a high leaf area index, green canopy.

2d. WRF Domain Configuration

WRF is a non-hydrostatic, fully compressible atmospheric model that is maintained by the National Center for Atmospheric Research (NCAR) for publically available source code distribution (Shamarock et al. 2008). This study used the advanced research dynamical core version of the WRF 3.4.1 model release with a double nested domain configuration as shown in Figure 6. The outer WRF model domain (labeled D1) had a horizontal grid spacing of 9 km and covered a large portion of the central United States. Domain 2 (D2) had a horizontal grid spacing of 3 km while domain 3 (D3) had a grid spacing of 1 km. The model domain configuration was centered on the SGP ARM Central Facility in north-central Oklahoma where the wind speed and energy flux observations used to validate the model performance are located. Given the absence of complex terrain in the innermost model domain, 1 km horizontal resolution was sufficient to resolve fine-scale horizontal flow in the domain of primary interest. A total of 50 terrain following vertical sigma levels were used for all of the WRF simulations not involving the Simplified Simple Biosphere Model (SSiB) LSM. Simulations using the SSiB LSM were run with only 41 vertical levels due to a model constraint that is discussed later. The 50 sigma level configuration resulted in a vertical resolution of roughly 20-30 meters in the lowest 200 meters of the atmosphere. The high vertical resolution in the surface layer was needed to resolve the complex, fine-scale vertical structure of the near-surface wind profile. The top of the model grid was at 50 hPa, which corresponds roughly to a height of 20 km above sea level.

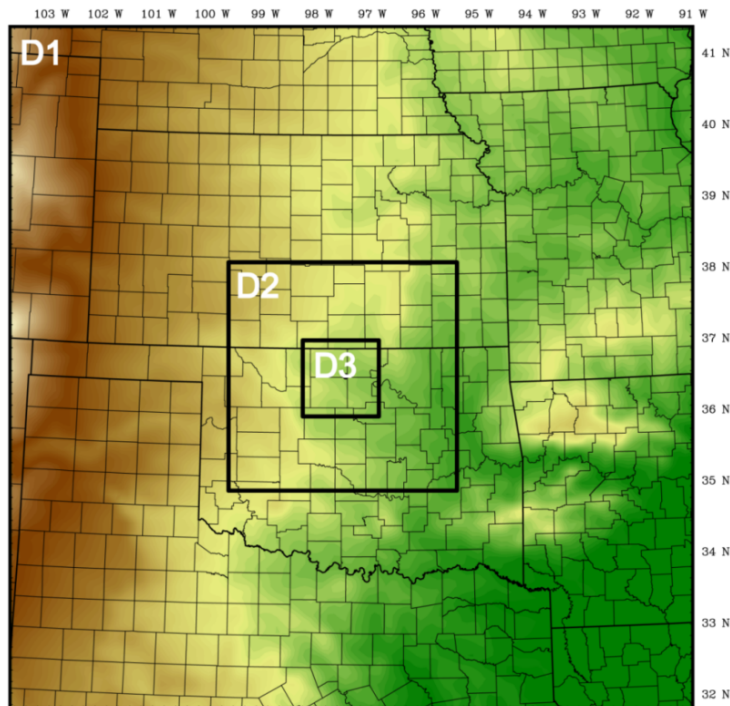


Figure 6 Elevation map showing the WRF model double nested domains used in this study. The inner domain, D3, is centered on the SGP ARM Central Facility in northern Oklahoma, USA. The horizontal grid resolution of domain 1 (labeled D1), domain 2 (D2), and domain 3 (D3) is 9, 3, and 1 km respectively.

2e. Ensemble Description

An ensemble consisting of the various LSMs available in the WRF Version 3.4.1 release was used to study the impact of LSM choice on simulating near surface wind profiles, with a focus on simulating hub-height wind speed (80 m AGL) and wind shear across heights equivalent to a modern turbine rotor disk (~40 – 120 m AGL). A total of 10 ensemble members were included in the study. All used Lin microphysics, the Kain-Fritsch cumulus scheme, and the CAM shortwave and longwave radiation model. Surface property input data came from the default, 33 USGS land use categories. The LSM, PBL, and surface layer schemes used for each of the ten members are listed in Table 2. The ensemble includes the 5-layer thermal diffusion (Blackadar 1978), Noah (Chen et al. 1996), RUC (Smirnova et al. 1997), Pleim-Xiu (PX; Pleim and Xiu 1995, Xiu and Pleim 2001), Noah Multi-Physics (Noah MP; Niu et al. 2011), and the SSiB LSM (Xue et al. 1991, Sun and Xue 2001). Various configurations of Noah MP and Pleim-Xiu were also run. A full description of the LSMs is provided in Section 2h.

Four of the ensemble members shown in Table 2 are used to represent the Noah MP, where the impact of the choice of soil moisture factor for stomatal resistance (opt_btr) and runoff / groundwater option (opt_run) was studied. Opt-btr was set as either option 2 (CLM model) or option 3 (SSiB model). Opt-run was set as either option 2 (TOPMODEL with equilibrium water table) or option 4 (BATS surface and subsurface runoff). All of the ensemble members used the Mellor-Yamada –Janjic (MYJ) PBL model and Monin-Obukhov surface layer scheme (sf_sfclay_physics = 2) except for members 9 and 10. The Asymmetric Convective Model (ACM2) PBL scheme (Pleim, 2007a,b) was used along with the PX LSM and Pleim-Xiu surface layer (option 7) in member 9 to represent a physics suite that was developed as a coupled configuration (Gilliam and Pleim, 2010) and is widely used in the atmospheric modeling community. As PX LSM could potentially perform best when coupled with the ACM2 PBL, it was included in the ensemble study to increase confidence in the PX LSM results. The PX LSM was also run with the MYJ PBL scheme in member 8 to allow direct comparison of model results with other members using the MYJ PBL. The Yonsei University (YSU) PBL model was used with the SSiB LSM and MM5 Monin-Obukhov surface layer scheme (option 1) in member 10 since YSU is the only PBL model that works with the SSiB LSM as of the WRF Version 3.4.1 release. It should also be noted that the SSiB LSM requires the first WRF full sigma level above the surface to be no greater than 0.982. This translates to a height of around 82 m for the first half sigma level as compared to a height of around 10 m for the other members. As a result, utilizing fine-scale vertical resolution near the surface is currently not possible with the SSiB LSM. This limited a direct comparison with the other members due to the differences in vertical resolution and also did not allow for the simulation of rotor-disk wind shear (40-140 m AGL) with member 10.

Table 2 Configuration of the WRF land surface model (LSM) based ensemble.

Member	LSM	PBL	Surface layer	Soil Layers	Noah opt_btr	Noah opt_run
1	Thermal	MYJ	MO	5		
2	Noah	MYJ	MO	4		
3	RUC	MYJ	MO	6		
4	Noah MP	MYJ	MO	4	2	2
5	Noah MP	MYJ	MO	4	2	4
6	Noah MP	MYJ	MO	4	3	2
7	Noah MP	MYJ	MO	4	3	4
8	Pleim-Xiu	MYJ	MO	2		
9	Pleim-Xiu	ACM2	Pleim-Xiu	2		
10	SSiB	YSU	MM5 MO	3		

2f. Input Data

Initial and lateral boundary conditions for the WRF simulations were provided by gridded analysis fields from the North American Model (NAM). The NAM analysis fields are available every 6-hours with a horizontal resolution of 12 km. Three-dimensional atmospheric data are provided on 40 pressure levels with a vertical resolution of 25 hPa in the NAM analysis data set. Meteorological variables provided by the NAM data include atmospheric pressure, geopotential height, horizontal wind components, air temperature, relative humidity, surface pressure, sea level pressure, and soil moisture and temperature at 4 sub-surface layers. NAM analysis data are available for public download from the National Climatic Data Center's (NCDC) data website (<http://nomads.ncdc.noaa.gov/data/namanl>).

2g. Four-Dimensional Data Assimilation

The WRF model has a four-dimensional data assimilation (FDDA) capability for both analysis (gridded) (Stauffer 1994) and observational nudging (Liu 2005). Analysis nudging involves using a Newtonian relaxation term to nudge the model integration towards an observed state; in this case represented by large-scale gridded analysis fields such as NAM data. Similarly, the WRF FDDA observational nudging model nudges localized model results towards individual weather observations. Using a FDDA nudging approach within an atmospheric model has been demonstrated to reduce error by constraining large-scale atmospheric flow towards analysis fields while allowing smaller scale atmospheric features to develop in fine-scale domains (Stauffer 1990, Lo 2008, Otte 2008, Salathe 2008, Bowden 2012). The degree to which the WRF numerical solution is nudged at each model grid point is a function of a user defined observation radius of influence, a relaxation time scale, and the distance to the observation.

The FDDA initialization methodology used for this study involved applying analysis nudging on model domains 2 and 3 to constrain the simulated large scale flow of the simulations. Model grid points within the planetary boundary layer were not nudged in either domain 2 or 3 to avoid conflicting with small

scale features that may develop. As an additional constraint to ensure that the model results in the lower atmosphere were not influenced by the nudging algorithm, the FDDA analysis option was turned off below model level 10, which corresponded to a height of roughly 600 m above the surface. The additional FDDA constraint was needed for extremely stable nighttime periods when boundary layer heights of less than 100 m were plausible. Observational nudging was not utilized in any of the WRF domains so that the model solution was free to evolve near the surface where results were later compared with atmospheric observations. A spin up period of four days was allowed before the WRF model output was compared with observations. A multi-day spin up period was necessary to ensure soil moisture and temperature initialization induced LSM biases had sufficient time to reach a reasonable balance with underlying atmospheric forcings. All of the model simulations were performed as a single continuous run with no restarts. Output from each of the ensemble forecast members was saved at 10 minute intervals for comparison with observations.

2h. Land Surface Models

Land surface models in WRF use atmospheric information from the surface layer scheme, radiative forcing information from the radiation scheme, and precipitation forcing from the microphysics and convective schemes, together with information about the land surface, to output heat and moisture fluxes. The Land Surface Models (LSMs) are described in this section in order of increasing complexity in regards to how they deal with thermal and moisture fluxes in the soil, and vegetation, root and canopy effects. The most basic of the LSMs used here is the thermal diffusion, or “slab”, model (Blackadar 1978). The thermal diffusion LSM calculates the surface heat fluxes from a 1-D equation and then assumes linear temperature profiles across 5 layers of soil. Soil moisture is fixed in time with a land use- and season-dependent constant value. There are also no explicit vegetation processes in the model.

The Rapid Update Cycle (RUC) model is largely focused on accurate characterization of the soil with up to 6 – 9 soil levels (or more possible) with the default parameterization reaching down to a soil depth of 300 cm. The physics of snow and phase change in soil are well characterized in this model (Smirnova et al. 1997). The model uses a basic approach to characterize evapotranspiration from the canopy based on work of Pan and Mahrt (1987), wherein the same equations used to calculate the transfer of heat and moisture from the soil are adjusted to estimate the fluxes from the canopy. Canopy evapotranspiration is scaled by potential evapotranspiration, canopy moisture content, rooting depth, and the soil moisture relative to the plant wilting point.

The Pleim-Xiu model (Pleim and Xiu 1995) builds off of the Noilhan and Planton (1989). This model has only 2 soil layers: a shallow layer (1 cm) at the surface interacting with the atmosphere, and a thick layer (99 cm) below interacting with the plants via the roots and the upper layer. Moisture fluxes are either from precipitation, dew formation, evaporation directly from the soil surface or vegetation, transpiration (ultimately controlled by soil moisture in the root zone via a stomatal resistance), or vertical movement within the soil. For the calculation of soil heat flux, only a soil heat capacity is used regardless of canopy cover or density.

The Simplified Simple Biosphere Model (SSiB) (Xue et al 1991, 1996) is a modification of the Simple Biosphere Scheme (SiB) of Sellers *et al.* (1986). SiB is a two-layer “big-leaf” model that scales leaf-level

processes to 2 representative ‘leaves:’ one each representing the canopy and ground cover. SSiB simplifies by reducing the model to a single-layered big-leaf model. SSiB requires 21 input parameters (reduced from 42) and additionally simplifies the diurnal variation of surface albedo for a given vegetation type and the effect of root zone soil moisture on stomatal resistance. Canopy evapotranspiration is controlled solely by physics.

The Noah model (Chen et al. 1996) is the most complex of the LSMs evaluated. Noah is the product of a large community of researchers combining previous iterations of multiple land surface models. The model has 4 soil layers reaching 2 m in which soil temperature and moisture are calculated. The shallow soil, ‘leaky’ bottom (precipitation that infiltrates to the bottom of the soil column is lost from the system) and simple snow/melt/thaw dynamics were seen as weaknesses. Vegetation, however, is well defined using monthly estimates of albedo and fraction of green vegetation cover. Evapotranspiration is modeled by the Ball-Berry equation, taking both the physics of water flow through the soil and plants as well as the physiology of photosynthesis into account.

Multi-parameterizations are available with the Noah model (Noah-MP) for users to tune the model in order to maximize complexity for dominant processes, while using simpler modeling approaches for other processes in order to conserve computational requirements (Niu et al. 2011). Noah-MP includes three options for modeling the response of transpiration to changes in site water status: via soil moisture directly as in Noah and 2 variations of the response of conductance to soil matric potential. Site hydrology can be modeled simply as in Noah with the ‘leaky’ bottom either with the simplified calculations of surface runoff or with more complex calculation of surface runoff based on the BATS model (Yang and Dickinson, 1996). Alternatively, the TOPMODEL-based (Niu et al. 2005) calculation of surface runoff and groundwater discharge is available with varying levels of complexity in modeling groundwater. Two variations were introduced for modeling radiation transfer through a canopy to give a user three levels of complexity from which to choose. Variations on the radiation scheme alter the manner that gaps in the canopy are simulated.

2i. Metrics

LSM performance in WRF was based on three metrics:

1. *Accuracy of surface energy flux exchanges.* The first metric was the difference between simulated H and LE fluxes and the flux tower measurements, including the root mean squared error (RMSE).
2. *Midday wind shear RMSE.* The second metric was RMSE between simulated wind shear and lidar-measured wind shear. The heights were based on a nearby turbine rotor diameter: 121 to 39 m AGL. Although the simulations did not always output these exact heights, differences were usually less than ± 10 m. Wind shear was calculated for all ensemble members except for SSiB (Ensemble 10).
3. *Hub-height RMSE.* The third metric was RMSE between simulated hub-height wind speed and the lidar-measured hub-height wind speed. Here, measured hub-height wind speed was based on the 80 m lidar wind speed. The simulated hub-height wind speed was based on the vertical level in WRF closest to 80 m, which was usually within ± 5 m.

3. Results

3a. June and July 2011 Case Study: Surface Meteorology

Although the June and July simulation periods are only separated by 20 days, the surface meteorological conditions were very different. These conditions included daily water availability, atmospheric water demand, and air temperatures measurements. From June 6-23, 2011, the average midday relative humidity (RH) was 37%, average midday air temperature (T_a) was 33.4 °C, while for the two week period in July, the midday values were warmer and drier: average midday RH was 29% and average midday T_a was 38.3 °C. In June, soil moisture was readily available to the plants (mean $\theta_v = 24\%$; wilting point $\theta_v = 10\%$) (Figure 7a) and the Bowen ratio on average was less than one (Figure 7c), so that slightly more energy was transferred to the atmosphere through evapotranspiration fluxes than through sensible heat during these two weeks. In July, however, the soil moisture content (mean $\theta_v = 6.4\%$) was well below the permanent wilting point (Figure 7b) and evapotranspiration or LE fluxes were very small. During this period, ten times more heat was transferred back to the atmosphere in the form of sensible heat than in the form of LE (mean $\beta = 10$) (Figure 7d). In Figure 7, daily soil moisture is based on an average of the lower and upper soil moisture measurement profile systems.

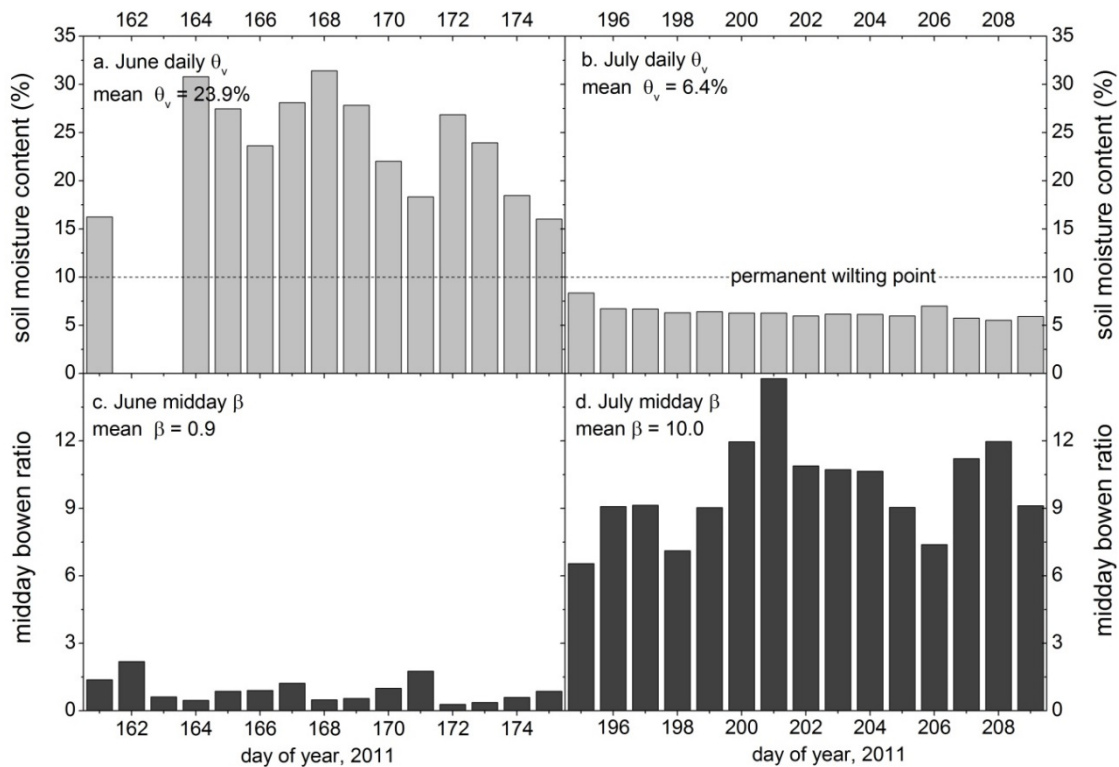


Figure 7 Mean daily soil moisture content measurements taken during the two week WRF periods in **(a)** June and **(b)** July 2011. Mean midday Bowen ratio measurements from the flux tower from the same days in **(c)** June and **(d)** July 2011. The permanent wilting point is for a well-drained Kirkland soil series.

3b. Land-Atmosphere Energy Exchange

In Objective (1), we examined the accuracy of each ensemble member to simulate midday sensible heat and latent energy fluxes. We also examined the ratio of the two or the Bowen ratio ($\beta = H/LE$). Accuracy was determined by the root mean square error between the ensemble member and observation. The measurement and simulated fluxes are shown in Table 3 and Figure 8 for each simulation period. The root mean square errors for sensible heat were highest for June 2011 (164 W m^{-2}) and lowest for Nov-Dec 2012 (25 W m^{-2}). For latent energy, the July 2011 period had the highest RMSE (150 W m^{-2}) while Nov-Dec 2012 had the lowest (33 W m^{-2}). This is partially a reflection of overall lower energy fluxes during the cool, dry autumn period. Overall, the most consistent and best performers used the Noah MP LSM, including E4 and E7, based on these members having the lowest RMSE for both sensible heat and latent energy fluxes. Surprisingly though, these ensembles were not always the best performers as some of the LSMs were both the best or worst performer depending on simulation period (Table 3).

During the June 2011 simulation E1 performed best in simulating sensible heat although overall it was one of the worst performing ensemble members. For example, E1 greatly overestimated midday latent energy fluxes during all simulation periods, by a factor of nearly 12:1 during the hot, dry July 2011 simulation, and underestimated the sensible heat flux during all simulation periods. E10 and E3 were also overall poor performers. Although E10 accurately predicted the LE flux in July 2011, it most often over predicted the sensible heat flux. The Bowen ratio was either very high (>15) or near infinity for both E3 and E10 during the 2011 simulation periods, as these LSM simulations predicted very low latent energy fluxes (Figure 8).

Table 3 Mean midday sensible heat and latent energy fluxes (W m^{-2}) for each ensemble member as well as the root mean square error (RMSE) for each simulated flux. The top performers are highlighted.

Midday sensible heat flux (W m^{-2})					
Ensemble Member	June 10-23, 2011	July 13-27, 2011	Nov 23-Dec 7, 2012	May 24-June 10, 2013	Mean RMSE
observations	152	330	111	not available	--
E1	86	67	51	103	160
E2	306	341	104	198	89
E3	413	445	98	324	164
E4	237	313	104	225	50
E5	263	304	99	242	66
E6	258	309	107	167	62
E7	239	316	106	173	51
E8	286	263	80	204	88
E9	221	263	75	196	59
E10	487	444	123	98	204
Ensemble flux mean (W m^{-2})	280	307	95	193	--
Ensemble flux error range (W m^{-2})	-66 (E1) to +335 (E10)	-263 (E1) to +115 (E3)	-60 (E1) to +12 (E10)	not available	--
Mean RMSE (W m^{-2})	164	103	25	not available	99

Midday latent energy flux (W m^{-2})					
Ensemble Member	June 10-23, 2011	July 13-27, 2011	Nov 23-Dec 7, 2012	May 24-June 10, 2013	Mean RMSE
observations	209	39	28	not available	--
E1	457	454	85	353	281
E2	180	80	52	291	32
E3	105	4	41	160	64
E4	198	55	14	216	14
E5	170	61	14	192	27
E6	169	46	13	287	25
E7	167	48	14	273	26
E8	200	202	70	236	97
E9	255	189	75	270	95
E10	58	35	75	450	91
Ensemble flux mean (W m^{-2})	196	117	45	273	--
Ensemble flux error range (W m^{-2})	-151 (E10) to +248 (E1)	-35 (E3) to +415 (E1)	-14 (E4, E5, E7) to +57 (E1)	not available	--
Mean RMSE (W m^{-2})	102	150	33	not available	75

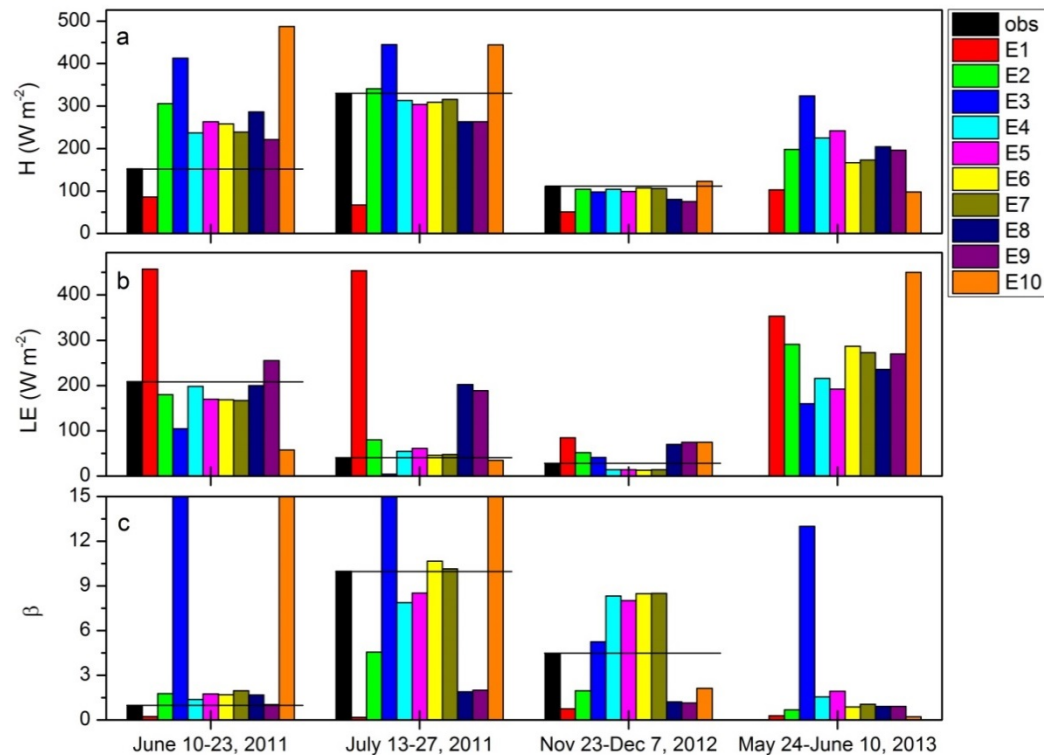


Figure 8 Mean midday (a) sensible heat, (b) latent energy, and (c) Bowen ratio values for each simulation period and ensemble member. The black lines indicate the measurement value and were drawn to ease comparison.

3c. Rotor-Disk Wind Shear

The two simulation periods in 2011 had very different energy partitioning values (Figure 8c) as the surface meteorology and canopy characteristics rapidly changed. We were curious to see if lidar-measured wind shear also had distinct differences, as we expected to see less wind shear during the July simulation period when buoyancy heat fluxes dominated the surface energy exchange. These strong buoyancy fluxes should produce a well-mixed daytime PBL. Observed differences in midday wind shear were smaller than expected although differences began to appear at 80 m AGL and above. Wind shear at heights equivalent to the upper half of a turbine rotor disk were slightly smaller in July (Figure 9a). In comparison, the Bowen ratio during this period was 10 times higher than in June. Figure 9 shows average conditions, normalized by hub-height wind speed, during the two periods, and some individual days had higher differences. Figure 9b shows wind shear during nighttime hours only. Here, case study distinctions are greater, especially after a height of 80 m is reached. July had much higher wind shear values in the upper half of the rotor disk which were likely caused by nocturnal low level jets.

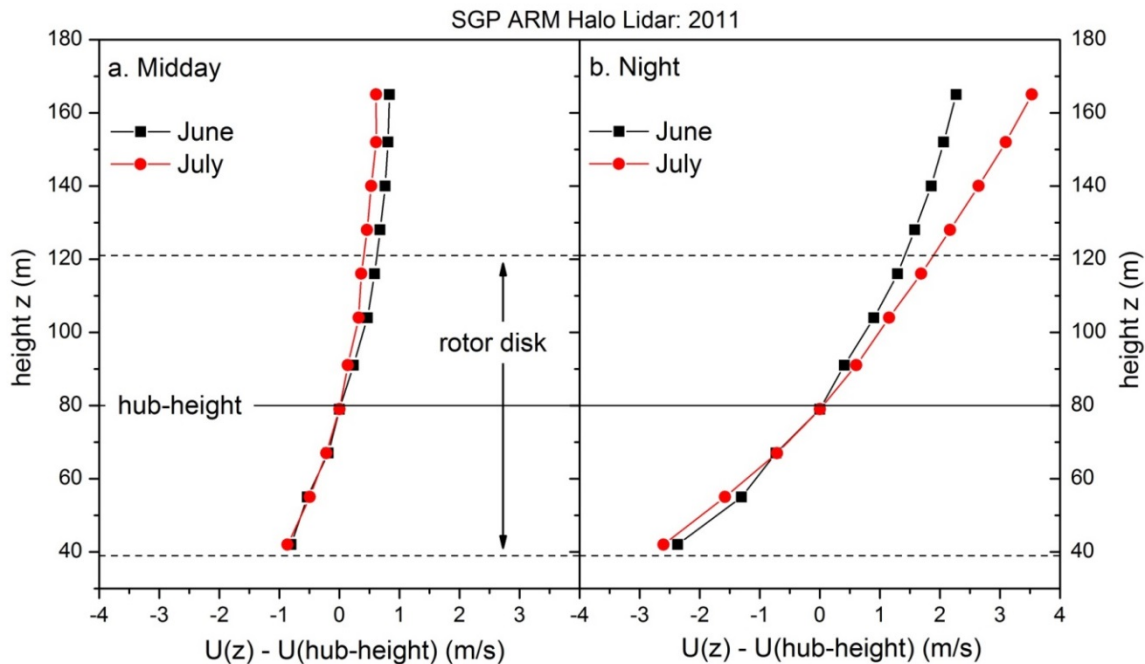


Figure 9 Wind speed profile measurements by time of day for the June and July simulation periods. The wind speeds are normalized by the 80 m U or hub-height wind speed.

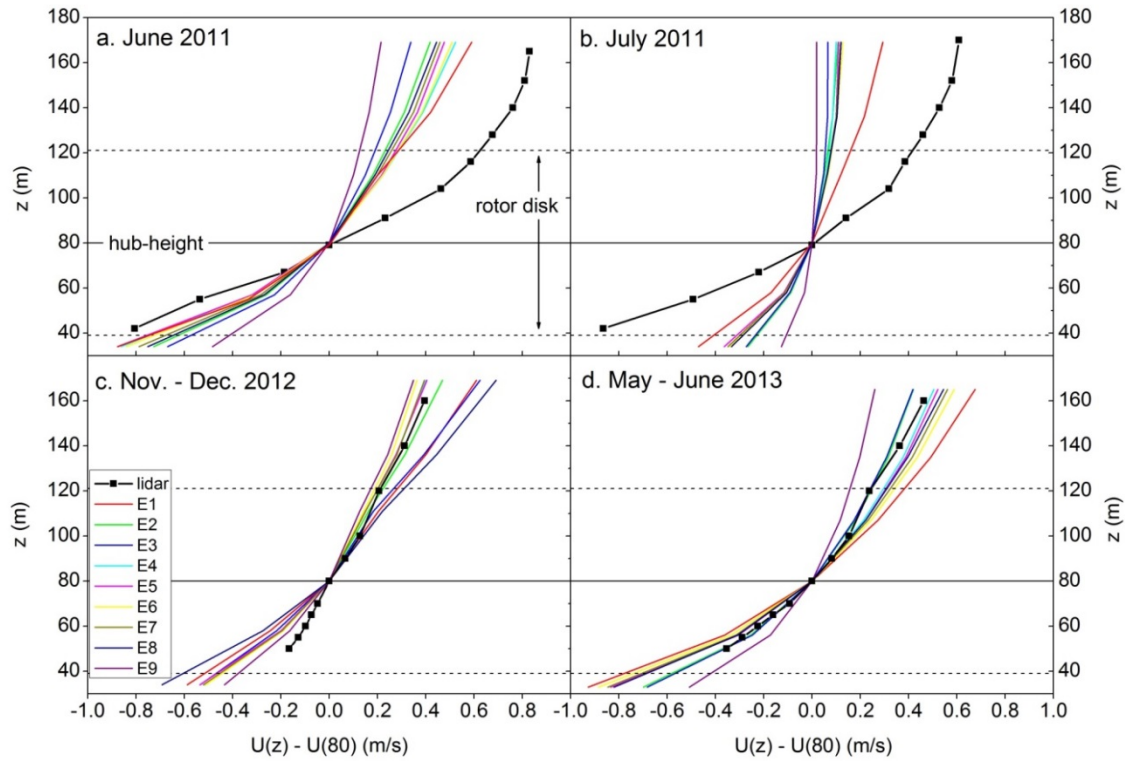


Figure 10 Midday wind speed profiles, normalized by hub-height (U at 80 m), during each two week simulation period for each ensemble member in comparison to the lidar observations shown in black.

Figure 10 shows the simulated midday wind shear profiles for all of the WRF periods, except E10, in comparison to the lidar observations. Midday wind shear error decreased from simulation 1-4, with no apparent explanation (Table 4). Interestingly, the WRF members predicted much less wind shear during July 2011 than June 2011 as we had expected to see if the observations. The WRF runs had the best overall performance during simulation 4 when conditions were warm, wet, and the ground was covered by a high LAI wheat canopy. Best performers over all were the Noah MP runs which also were the most accurate at predicting the surface energy flux exchange (see Table 3 and Figure 8). RMSE are plotted for both daytime and all hours in Figure 11. Overall, the RMSE for wind shear was higher for all hours of the day than for daytime hours only. In some cases, this difference was a factor of two or more. In particular the thermal diffusion, E1, and Pleim-Xu, E8 and E9, models did particularly poorly at night when predicting wind shear.

Table 4 Midday RMSE (m/s) for rotor-disk wind shear (~120 m to ~40 m) when hub-height wind speed was between cut-in and rated wind speed ($3.5 < U < 12.5$ m/s). Top performers are highlighted.

Midday wind shear RMSE (m/s)					
Ensemble Member	June 10-23, 2011	July 13-27, 2011	Nov 23-Dec 7, 2012	May 24-June 10, 2013	Mean RMSE \pm SD
E1	1.40	1.42	1.10	0.80	1.18 ± 0.30
E2	1.42	1.29	0.87	0.64	1.06 ± 0.36
E3	1.44	1.35	1.53	0.66	1.25 ± 0.40
E4	1.47	1.30	0.69	0.63	1.02 ± 0.43
E5	1.39	1.28	0.73	0.72	1.03 ± 0.36
E6	1.40	1.30	0.70	0.83	1.06 ± 0.34
E7	1.35	1.29	0.67	0.78	1.02 ± 0.35
E8	1.40	1.35	1.01	0.74	1.13 ± 0.31
E9	1.89	1.49	0.81	0.63	1.21 ± 0.59
E10	--	--	--	--	--
RMSE range	1.35 to 1.89	1.28 to 1.49	0.67 to 1.53	0.63 to 0.83	1.02 to 1.25
Mean RMSE \pm SD	1.46 ± 0.17	1.34 ± 0.07	0.90 ± 0.28	0.71 ± 0.08	1.10 ± 0.32

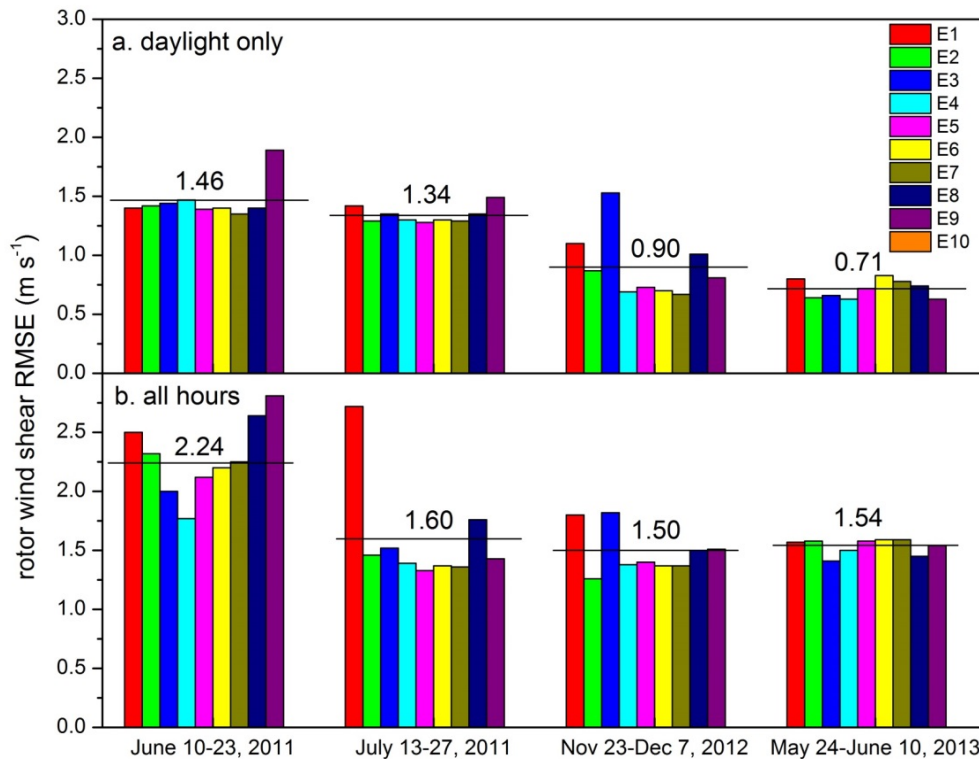


Figure 11 Root mean square errors for wind shear for each ensemble member and period during (a) daylight hours only, and (b) all hours of the day. E10 is not shown because this LSM has only one vertical height below 200 m AGL. The black lines indicate the mean RMSE for each simulation period.

3d. Hub-Height Wind Speed

It is standard practice for the wind industry to forecast hub-height wind speed as operational power curve models are usually a function of this wind speed alone. Thus the industry relies heavily on hub-height speed to forecast farm power generation. In this section we compare simulated wind speeds at ~80 m AGL from each ensemble member to the lidar observations, first for daylight hours only (Figure 12a, Table 5) and then for all hours of the day (Table 6, Figure 12b). RMSEs for daytime hub-height wind speed were on average 2-3 times greater than midday wind shear RMSEs (Section 3c). There was also more variability in the RMSEs indicating a larger range of errors for this forecasted wind speed than wind shear. Best performing LSMs depended on the simulation period and included E1, E2, E6, E8, E9 and E10; the thermal diffusion, Noah, Noah MP, Pleim-Xiu and SSiB models, respectively. Although a top performer during one of the simulation periods, the thermal diffusion model was extremely inconsistent and was also one of the worst performing models overall. The most consistently accurate LSM during all four simulation periods was E9, the Pleim-Xiu model run with the ACM2 PBL scheme. In second place was Pleim-Xiu with the MYJ PBL scheme. The worst performer was the RUC model (Table 5). Simulations of nighttime hub-height wind speed were slightly worse than daytime simulated wind speed, particularly for E10, the SSiB model (Figure 12b).

Table 5 RMSE (m/s) for hub-height wind speed (~80 m, m/s) during daylight hours only for wind speeds between cut-in and rated wind speed ($3.5 < U < 12.5$ m/s). Top performers are highlighted.

Daytime hub-height wind speed RMSE (m/s)					
Ensemble Member	June 10-23, 2011	July 13-27, 2011	Nov 23-Dec 7, 2012	May 24-June 10, 2013	Mean RMSE \pm SD
E1	3.06	1.97	2.93	2.29	2.56 ± 0.52
E2	2.88	2.05	2.25	2.28	2.36 ± 0.35
E3	2.71	2.39	3.06	2.62	2.69 ± 0.28
E4	3.22	1.92	1.78	2.91	2.45 ± 0.72
E5	2.87	2.02	1.81	3.09	2.44 ± 0.63
E6	3.09	1.90	1.72	3.13	2.46 ± 0.75
E7	3.03	2.13	1.78	2.73	2.42 ± 0.57
E8	2.77	1.84	2.43	2.27	2.32 ± 0.39
E9	2.40	1.90	2.31	2.35	2.24 ± 0.23
E10	3.35	1.86	1.70	2.52	2.36 ± 0.75
RMSE range	2.40 to 3.35	1.84 to 2.39	1.70 to 3.06	2.27 to 3.13	2.24 to 2.69
Mean RMSE \pm SD	2.94 ± 0.28	2.00 ± 0.16	2.18 ± 0.50	2.62 ± 0.33	2.43 ± 0.48

Table 6 RMSE (m/s) for hub-height wind speed (~ 80 m, m/s) during all hours of the day for wind speeds between cut-in and rated wind speed ($3.5 < U < 12.5$ m/s). Top performers are highlighted.

All hours hub-height wind speed RMSE (m/s)					
Ensemble Member	June 10-23, 2011	July 13-27, 2011	Nov 23-Dec 7, 2012	May 24-June 10, 2013	Mean RMSE \pm SD
E1	3.55	2.09	3.38	2.65	2.92 ± 0.68
E2	3.22	2.11	2.35	2.96	2.66 ± 0.52
E3	2.90	2.16	3.93	3.07	3.02 ± 0.73
E4	3.15	1.84	1.87	3.04	2.48 ± 0.72
E5	2.92	1.89	1.87	3.10	2.44 ± 0.66
E6	3.14	1.88	1.86	3.02	2.48 ± 0.70
E7	3.14	1.88	1.88	3.00	2.48 ± 0.69
E8	3.16	1.88	2.71	2.74	2.62 ± 0.54
E9	2.97	1.59	2.69	2.93	2.55 ± 0.65
E10	5.64	2.18	2.53	2.92	3.32 ± 1.58
RMSE range	2.90 to 5.64	1.59 to 2.16	1.86 to 3.93	2.65 to 3.10	2.44 to 3.32
Mean RMSE \pm SD	3.38 ± 0.82	1.95 ± 0.18	2.51 ± 0.71	2.94 ± 0.14	2.70 ± 0.75

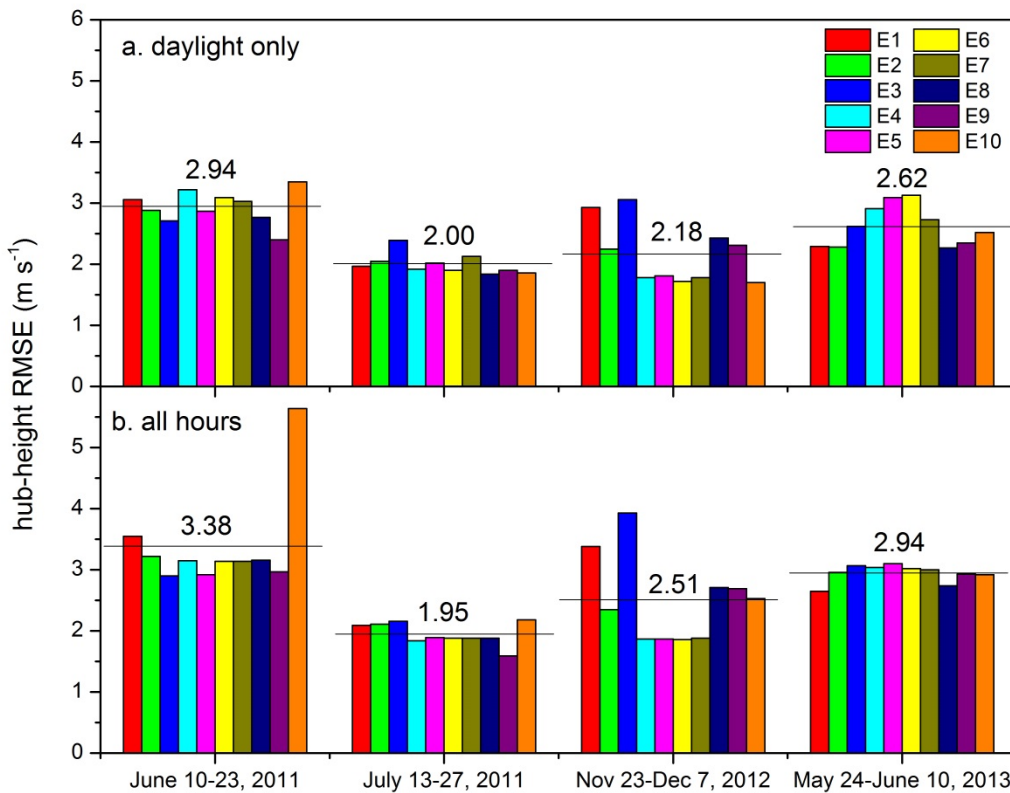


Figure 12 Root mean square errors for hub-height wind speed for each ensemble member and ensemble period during **(a)** daylight hours only, and **(b)** all hours of the day. Hub-height wind speed is approximately 80 m AGL. The black lines represent mean RMSE for each simulation period.

Based on the hub-height metric, the best and worst performing LSMs are plotted in Figures 13-16. Performance was based on the average RMSE for all hours of the day during the simulation period, excluding the first four spin-up days (Table 6). In these figures hub-height wind speed is plotted for the best and worst performing ensemble members in comparison to the lidar measurements.

For the June 2011 case, the observed synoptic weather pattern was simulated by with all ensemble members although the exact wind speed magnitudes were most accurately simulated with RUC (E3) and least accurate with SSiB (E10) (Figure 13). SSiB was particularly poorly with simulating nighttime hub-height wind speed as the predicted nighttime magnitudes were much higher than actual conditions and sometimes exceeded the observed value by 10 m/s or more. The average RMSE for all ensemble members was 3.38 ± 0.82 m/s; the highest of all four simulation periods.

During the July 2011 case, hub-height wind speeds were strongly forced by local surface energy exchange as the diurnal pattern is the dominant source of wind speed variability up to day 205, versus the synoptic forcing pattern observed in Figure 13. During this period all of the ensemble models had difficulty with simulating the magnitude of the wind speed troughs during midday, as nearly all models over predicted the daytime wind minima. Even with this error, the average RMSE for all ensemble members was 1.95 ± 0.18 m/s and nearly 1.5 m/s less than the June simulation. Reduced error was also evident for simulated daytime wind speeds only as compared to the June runs. The best performer during this period was Pleim-Xiu run with the ACM2 PBL scheme (E9) and the worst performer was the same as during the June period and was the SSiB model (E10).

During the Nov-Dec 2012 simulation the average RMSE was 2.51 ± 0.71 m/s. The best performing ensemble member was Noah MP (E6) run with options 3 (SSiB model for stomatal conductance) and 2 (TOPMODEL with equilibrium water table) and the worst performer was the RUC model (E3). Note that the RUC model was the best performer in the June 2011 case. In the autumn, the RUC model did particularly poorly with simulating days 333-337, which caused it to have a high RMSE. During these days the average wind speed was around 10 m/s while the RUC model predicted an average wind speed closer to 2 m/s (Figure 15). Forcing during Simulation 3 was a mixture of synoptic and locally-driven, diurnal flow.

During May-June 2013 the best performing model was the thermal diffusion model (E1) and the worst performing model was the Noah MP run with options 2 (CLM model for stomatal conductance) and 4 (BATS model surface and subsurface runoff) (E5) (Figure 16). E5 frequently “overshot” the wind maxima or minima leading to a higher RMSE. The May-June period had the second highest mean RMSE (2.94 ± 0.14 m/s) of the four study cases. Weather forcing was strongly synoptic during the first half of Simulation 4 and became more diurnal or locally driven during the second half of the period.

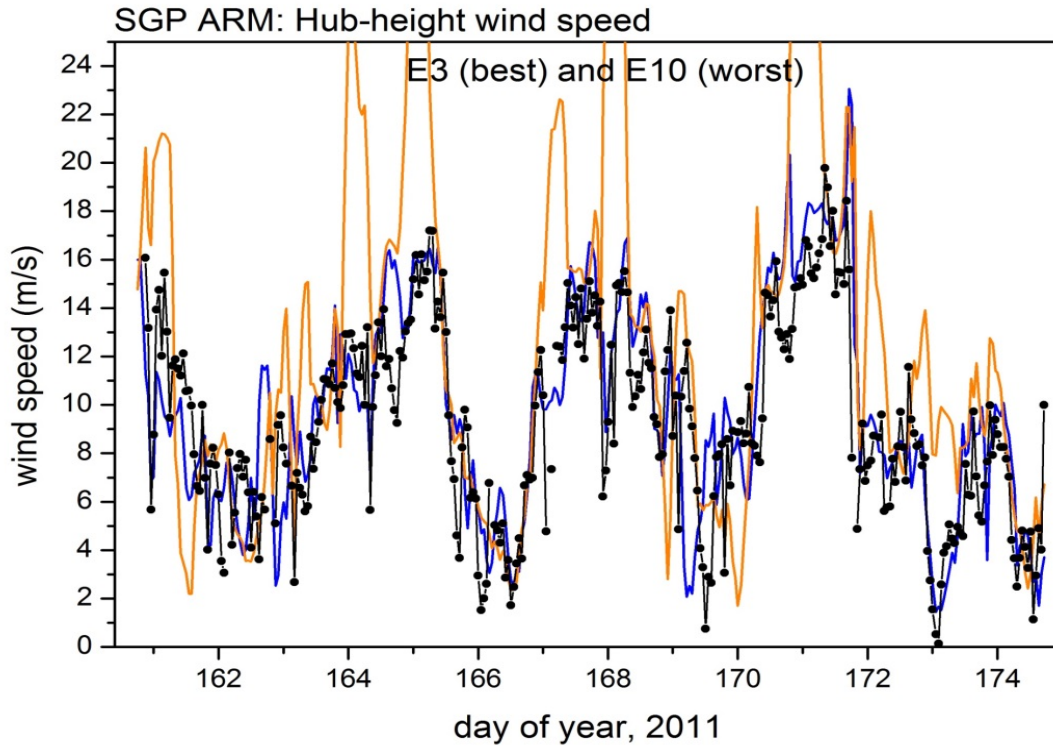


Figure 13 Simulation 1: Hourly hub-height wind speed measured by the lidar (observations, black circles) as compared to the best (E3, blue line) and worst (E10, orange line) performing ensemble members. Simulation period is from June 10-23, 2011 and time is in Central Standard Time.

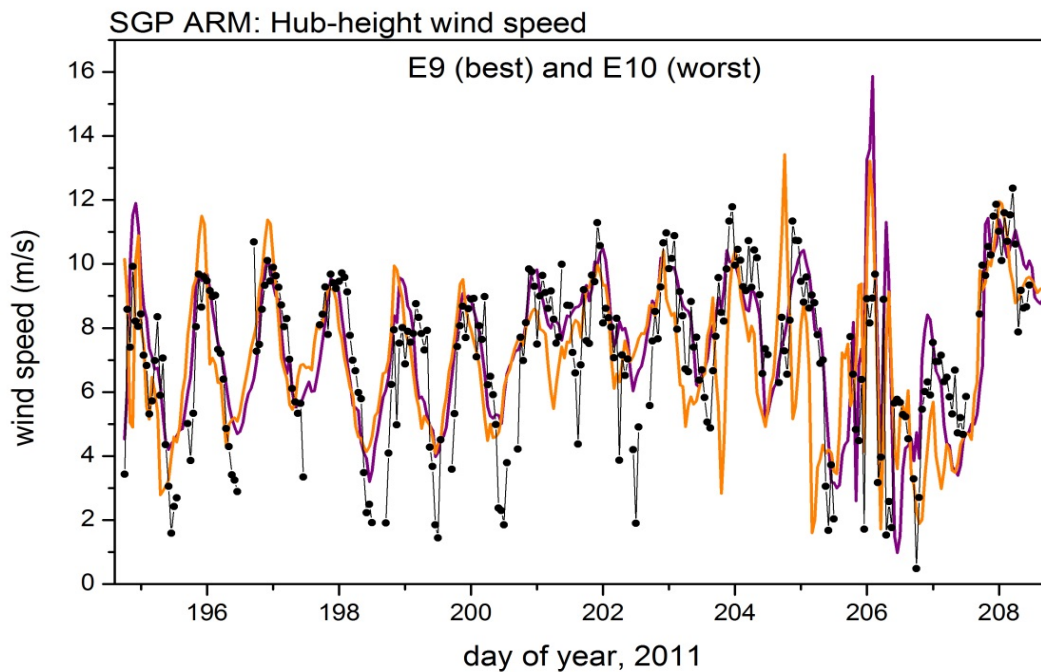


Figure 14 Simulation 2: Hourly hub-height wind speed measured by the lidar (observations, black circles) as compared to the best (E9, purple line) and worst (E10, orange line) performing ensemble members. Simulation period is from July 13-27, 2011 and time is in Central Standard Time.

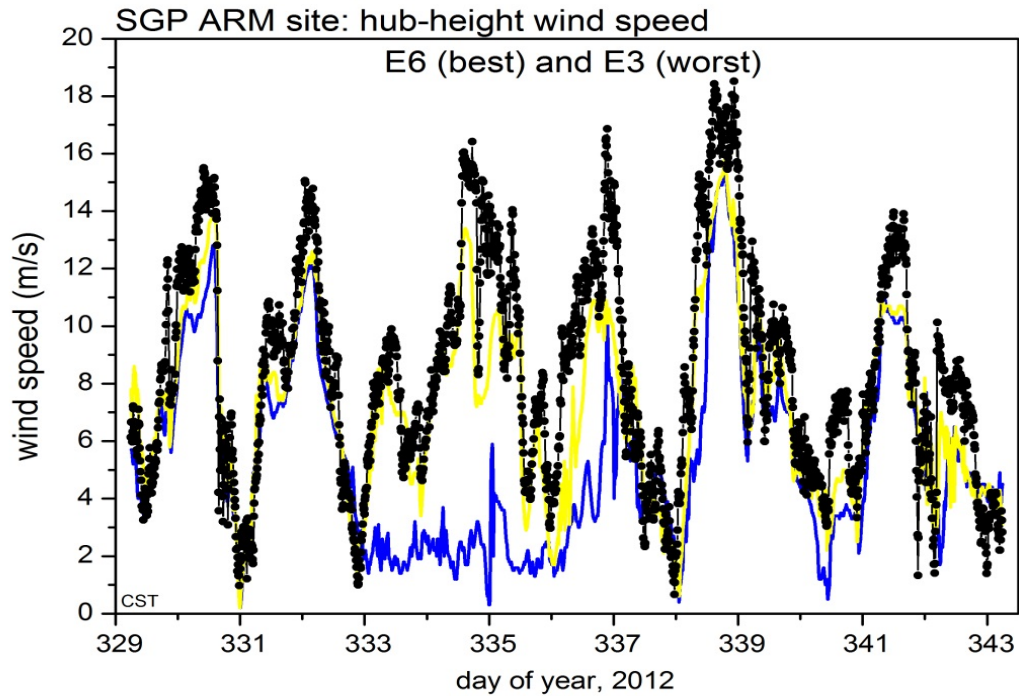


Figure 15 Simulation 3: 10-minute hub-height wind speed lidar measurements (observations, black circles) as compared to the best (E6, yellow line) and worst (E3, blue line) performing ensemble members. Simulation period is from November 23 – December 7, 2012 and time is in CST.

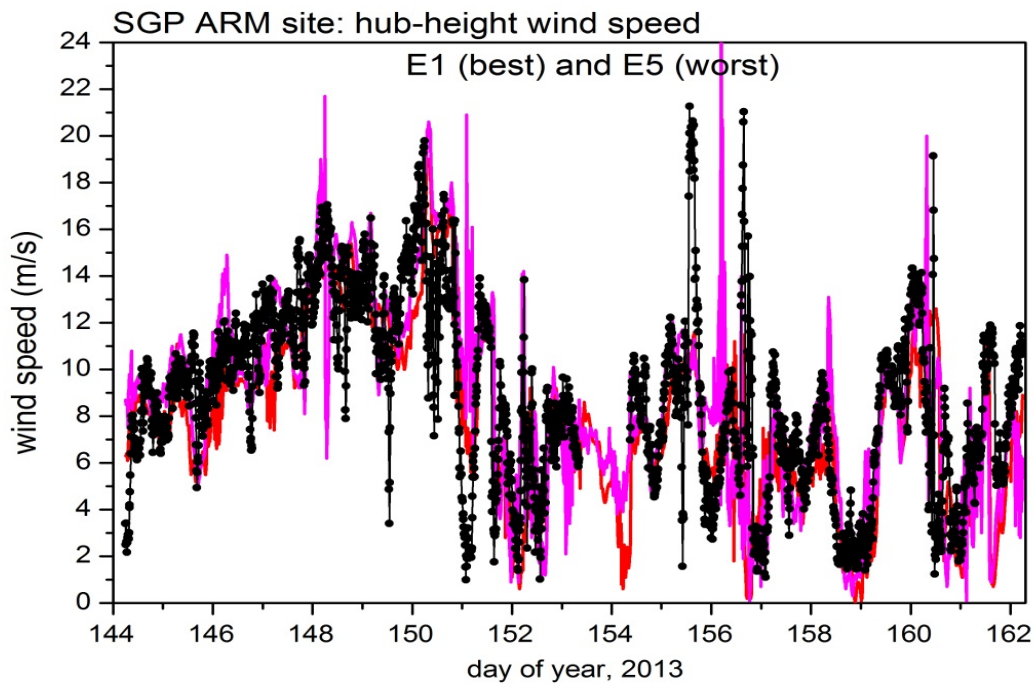


Figure 16 Simulation 4: 10-minute hub-height wind speed lidar measurements (observations, black circles) as compared to the best (E1, red line) and worst (E5, pink line) performing ensemble members. Simulation period is from May 24 – June 10, 2013 and time is in CST.

4. Discussion

Oklahoma ranked sixth in the nation for total wind power capacity at the end of 2012 (3,134 MW) and fourth in the nation for new capacity added in 2012 (1,127 MW) (Riser et al. 2013). While the landscape is flat and topographic flows are not as important as in other parts of the nation with complex terrain, wind forecasting in the Southern Great Plains continues to have other challenges. We conducted this study to investigate if observed wind shear and hub-height wind speed in this region are sensitive to land surface energy exchange and whether the choice of land surface model in WRF significantly changes the simulated wind fields at rotor-disk height. Earlier studies have shown that fine scale spatial variability, climate gradients, and climate anomalies (e.g., moderate to extreme drought) make the prediction and simulation of surface energy exchanges especially difficult in this part of the country (Fischer et al. 2007). The role of surface energy exchange and LSM choice in WRF for simulating wind power-relevant wind fields has until now received far less attention.

Land surface models become increasingly important in numerical weather prediction (e.g., WRF) models when: (1) more complex PBL schemes are used as they are more sensitive to surface fluxes, and (2) when the grid-spacing is on the order of a km and surface variability in albedo, soil moisture and land use become more important (Chen 2007). Here, we found that the proper choice of land surface model for simulated hub-height wind speed led to an overall improvement of 14% when the best performing model, Noah, was run versus the worst performing model, SSiB. Compared to the three worst performing LSMs, the forecasting improvement by Noah was nearly 8%. These improvements are significant considering that a forecasting improvement of 10-20% could result in hundreds of millions of dollars (\$140-260M) in annual operating cost savings for the wind industry (Lew et al. 2011).

The worst performing LSMs, thermal diffusion, RUC, and SSiB, for simulating wind speed also did a poor job at simulating the surface energy exchange and midday wind shear. The thermal diffusion model greatly overestimated the latent energy flux and underestimated the sensible heat flux while the RUC model performed the opposite way and overestimated H and underestimated LE. Both RUC and thermal diffusion models had high wind shear errors. These two LSMs were the two simplest used in this study and their poor performance with simulating the land surface energy budget is likely a direct result. The surface exchange errors carried over to wind shear and hub-height wind speed. The third poor performer, SSiB, is a relatively complex LSM but this model might have had problems with simulating wind speeds at height relevant to a turbine rotor disk because the model only includes one grid point below 200 m. This is seen by the fact that SSiB performed better in the likely strong Convective Boundary Layer of July 2011 where mixed conditions minimize the limitations of the coarse vertical resolution and poorest at night when vertical resolution closest to the ground is important for capturing wind shear.

The most complex LSM, Noah and Noah MP, performed the best overall for predicting both the surface energy exchange and the wind fields. Noah includes a vegetation evapotranspiration calculation where water lost from the canopy is parameterized in terms of “canopy resistances” to the “potential” water flux. This equation also includes a photosynthetically green fraction which is based on 5-year average, monthly Normalized Difference Vegetation Index (NDVI) values. The resistance calculation should

reproduce stressed vegetation and reduced water loss under hot and dry conditions and we saw this for the Noah runs in July 2011. Noah's simulations of the latent energy flux were more accurate than nearly all of the other LSMs during this very hot and dry period.

We saw the greatest hub-height wind speed errors when the canopy was actively photosynthesizing and the Bowen ratio was close to 1. This suggests that accurate simulations of the surface energy budget are particularly important for accurately modeling hub-height wind speed. Errors in hub-height wind speed were smaller when H and LE fluxes were either relatively small, such as during the autumn of 2012, or when the canopy was senesced and dry and LE fluxes were very small as during the hot July 2011 simulation period. Wind shear error decreased steadily from Simulation 1 through Simulation 4. The reason for this is unclear as Simulation 4 was during a period of strong land surface exchange and high evapotranspiration fluxes, similarly to Simulation 1.

A portion of the simulation errors could be due to the mismatch between spatial and temporal scales. We ran WRF down to 1 km resolution while the flux tower and lidar instruments had a footprint closer to 200 m. Flow could have significant variability even across these scales as land use changes and soil moisture gradients change rapidly in this area. WRF was run with a land use category defined as Dry Cropland and Pastureland while in reality the canopy type varied across space and time during the simulation periods. The LSM used here with the finest temporally adjusted land cover input was Noah which used a monthly time step of long-term NDVI averages. Further improvements to simulating the wind fields in this area likely could be done with using land use data with increased spatial and temporal resolution. For example, Riley et al. (2009) found that the Southern Great Plains region requires the characterization of vegetation properties in LSMs at a realistic spatial scale for accurate estimates of surface energy fluxes. The SGP region may need LSMs with the capability of having accurate specification of surface characteristics on a sub-grid scale to capture small scale variability, including tiling and harvest events.

A second source of error, during nighttime periods, could be WRF's difficulty with modeling nocturnal low-level jets (LLJ). Nighttime wind speed and wind shear errors in our simulations were larger than those during the daytime and nighttime errors were particularly high for SSiB which had very coarse vertical resolution near the ground. LLJs are frequent in this area during the spring and summer months. In fact, LLJs were frequently observed during a multi-lidar campaign which took place during our fourth study period, May-June 2013. These LLJs ranged from 250-1000 m AGL (Tim Bonin, personal communication). The occurrence of LLJs may be the cause for higher nighttime error during the first simulation period in June 2011. Higher nighttime modeling errors have also been found for the region by other groups (e.g., Storm and Basu 2010).

For hour ahead or day ahead operational forecasts, computational expense or "run time" is important in addition to forecast accuracy. Because we did not use a fixed time step, to get around deep convection errors, and instead used a WRF adaptive time step, firm conclusions based on computational costs are difficult to make. The CPU cost differences between ensemble members were not directly related to the "cost" of the physics choice. Given this caveat, some distinctions can be drawn. The SSiB model (Member 10) ran the fastest of all the simulations due to the fact that it had fewer vertical levels. Noah

MP Member 7 took the longest to run for all simulations although reasons for this are difficult to make given we used the adaptive time step.

This study was conducted as an evaluation of running each LSM in WRF in “operational mode” as a wind farm operator might do, without fine-tuning each model to ensure the best performance. For example, some of the LSMs investigated here have options for remote sensing surface data assimilation (e.g, RUC, Noah) which may have enhanced model performance. With this caveat and with the understanding that we ran each LSM in “standard mode”, we showed here that the choice of LSM significantly led to an improvement in simulating accurate hub-height wind speed. This is a significant finding as any improvement in short-term forecasting should drive down the cost of wind energy and further wind as a leading form of low cost renewable energy in the coming decade (EERE 2012).

5. Conclusion

The best and worst performing members found in this study were chosen by their overall ability to accurately simulate surface energy exchange, midday wind shear at heights equivalent to a turbine rotor disk, and hub-height wind speed (Table 7).

Table 7 Best and worst performing ensemble members for each metric.

	Metric 1: surface fluxes	Metric 2: wind shear	Metric 3: hub-height U
Best performers	Noah MP (E4)	Noah MP (E4, E5, E7)	Noah MP (E4-E7)
Worst performers	Thermal diffusion (E1)	RUC (E3) , Pleim-Xiu with ACM2 (E9)	SSiB (E10), RUC (E3)

Based on this study the following conclusions can be made:

1. Hub-height wind speed and wind shear at heights relevant to the wind industry are sensitive to land surface-atmospheric exchange processes, especially during periods of active canopy growth, as shown by lidar and flux tower observations.
2. Simulations of rotor-disk wind fields are sensitive to getting the surface exchange processes correct in the LSM and choice of LSM in WRF led to a significant (~10%) improvement in simulating hub-height wind speed at the SGP ARM site.
3. Noah MP ensemble members were the overall best performers in all metrics. This strongly suggests that Noah MP is one of the best LSMs in WRF for *wind energy forecasting* for this region.

6. Warranted Future Work

Four areas stand out as candidates for future investigation as these topics were beyond the scope of this work:

1. There was an inconsistency of LSM performance across the study periods which warrants more attention. The lack of consistent superior performance by one LSM suggests that a LSM might be right for the wrong reason (e.g. model BIAS cancelled out by inaccurate soil moisture / surface conditions). In numerical modeling, sometimes two wrongs do make a right.
2. An economic benefit analysis would be required to see if any increases in LSM accuracy are worth additional CPU expenses. This requires cross comparing WRF CPU time with LSM model wind speed / power prediction performance to understand if running a specific LSM makes economic sense. For example, it may take 10% longer to run a specific LSM but it may provide 15% more accurate wind speed predictions. Having this information is essential for recommending a best performing LSM to the wind energy community.
3. The addition of near real-time data assimilation from satellite products such as MODIS NDVI, soil moisture products, and LAI, needs to be investigated. High-resolution satellite data would improve the realism of surface properties in the coupled WRF-LSM model. For example, crops at SGP ARM site are harvested at different times of the year depending on changing weather conditions, the crop field is tilled further changing albedo and surface soil moisture, crops quickly senesced over a period of a couple of weeks, but this information is not in the standard versions of WRF. For example, WRF was reinitialized with new soil moisture conditions between the June and July 2011 simulations but it used the *same* land use based surface characteristics even though in reality the canopy senesced, the crop was harvested, and the field was tilled during these four weeks in 2011.
4. How universal are the findings for other parts of the country? Concurrent lidar and energy flux data were collected in 2012-2013 at LLNL's Site 300 Experimental Site in the Altamont Wind Resource Region of California. These data need to be evaluated as the terrain is much more complex than at the SGP ARM site. A follow up study is also planned for WFIP 2.0 for a wind energy region in complex terrain.

7. Acknowledgements

This work was funded by the Department of Energy, Wind and Water Power Technologies Office under agreement, WBS 02.07.00.01. The authors would like to thank the staff at the DOE SGP ARM Central Facility for their help with maintaining the Wind Cube lidar system and their fantastic support throughout our study. We'd also like to thank Marc Fischer and Margaret Torn at LBNL with their support with the surface flux tower data. LLNL is operated by Lawrence Livermore National Security, LLC, for the DOE, National Nuclear Security Administration under contract DE-AC52-07NA27344.

8. References

- Betts, A.K., Ball, J.H., Beljaars, A.C.M., Miller, M.J., Viterbo, P.A. (1996) The land surface-atmosphere interaction: A review based on observational and global model perspectives. *J. Geophysical Research* 101: 7209-7225
- Blackadar, A.K. (1978) Modeling pollutant transfer during daytime convection. Preprints, *Fourth Symposium on Atmospheric Turbulence, Diffusion and Air Quality*, Reno, Nevada, American Meteorological Society, 443-447
- Bowden, J.H., Otte, T.L., Nolte, C.G., Otte, M.J. (2012) Examining Interior Grid Nudging Techniques Using Two-Way Nesting in the WRF Model for Regional Climate Modeling. *J. Climate* 25: 2805–2823
- Case, J.L., LaFontaine, F.J., Kumar, S.V., Peters-Lidard, C.D. (2012) Using the NASA-Unified WRF to assess the impacts of real-time vegetation on simulations of severe weather. Preprints, 13th Annual WRF User's Workshop, Boulder, CO, p69
- Chen, F., Mitchell, K., Schaake, J., Xue, Y., Pan, H.-L., Koren, V., Duan, Q.Y., Ek, M., Betts, A. (1996) Modeling of land surface evaporation by four schemes and comparison with FIFE observations. *J. Geophysical Research* 101: 7251-7268
- Desai et al. 2006
- Draxl, C. Hahmann, A.N., Pena, A., Giebel, G. (2012) Evaluating winds and vertical wind shear from Weather Research and Forecasting model forecasts using seven planetary boundary layer schemes. *Wind Energy*, in press, DOI: 10.1002/we.1555
- Gilliam, R.C., Pleim, J.E. (2010) Performance Assessment of New Land Surface and Planetary Boundary Layer Physics in the WRF-ARW. *J. Appl. Meteor. Climatol.* 49: 760–774
- LeMone, M.A., Grossman, R.L., Coulter, R.L., Wesley, M.L., Klazura, G.E., Poulos, G.S., Blumen, W., Lundquist, J.K., Cuenca, R.H., Kelly, S.F., Brandes, E.A., Oncley, S.P., McMillen, R.T., Hicks, B.B. (2000) Land-atmosphere interaction research and opportunities in the Walnut River watershed in southeast Kansas: CASES and ABLE. *Bull. Amer. Meteor. Soc.* 81: 757–779
- LeMone, M.A., Chen, F., Alfieri, J.G., Tewari, M., Geerts, B., Miao, Q., Grossman, R.L., Coulter, R.L. (2007) Influence of land cover and soil moisture on the horizontal distribution of sensible and latent heat fluxes in Southeast Kansas during IHOP_2 and CASES-97. *J. Hydrometeorology* 8: 68-87
- Lew, D., Milligan, M., Jordan, G., Piwko, R. (2011) Value of Wind Power Forecasting, NREL Report No. CP-5500-50814, 13p
- Lo, J. C. F., Yang, Z.L., Pielke, R.A., Sr. (2008) Assessment of three dynamical climate downscaling methods using the Weather Research and Forecasting (WRF) model. *J. Geophys. Res.* 113: doi:10.1029/2007JD009216
- Liu, Y., A. Bourgeois, T. Warner, S. Swerdlin, Hacker, J. (2005) An implementation of obs-nudging-based FDDA into WRF for supporting ATEC test operations. WRF User Workshop, Boulder, Colorado, Paper 10.7

Maxwell, R.M., Chow, F.K., Kollet, S.J. (2007) The groundwater-land surface-atmosphere connection: soil moisture effects on the atmospheric boundary layer in fully-coupled simulations. *Advances in Water Resources* 30: 2447-2466.

Miller, J., Barlage, M., Zeng, X., Wei, H., Mitchell, K., Tarpley, D. (2006) Sensitivity of the NCEP/Noah land surface model to the MODIS green vegetation fraction data set. *Geophys. Res. Lett.* 33: doi:10.1029/2006GL026636

Newsom, R. K. (2012) Doppler Lidar Handbook. Technical report # DOE/SC-ARM-TR-101, 23 p.

Niu, G.-Y., Yang, Z.-L., Dickinson, R.E., Gulden, L.E. (2005) A simple TOPMODEL-based runoff parameterization (SIMTOP) for use in global climate models, *J. Geophys. Res.* 110: doi:10.1029/2005JD006111

Niu, G.-Y., Yang, Z.-L., Mitchell, K.E., Chen, F., Ek, M.B., Barlage, M., Kumar, A., Manning, K., Niyogi, D., Rosero, E., Tewari, M., Xia, Y (2011) The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. *J. Geophys. Res.* 116: doi: 10.1029/2010JD015139

Noilhan J., Planton, S. (1989) A simple parameterization of land surface processes for meteorological models. *Mon. Wea. Rev.* 117, 536-549

Otte, T. L. (2008) The Impact of Nudging in the Meteorological Model for Retrospective Air Quality Simulations. Part I: Evaluation against National Observation Networks. *J. Appl. Meteor. Climatol.* 47: 1853–1867

Pan, H.-L., Mahrt, L. (1987) Interaction between soil hydrology and boundary layer development, *Boundary Layer Meteorol.* 38: 185-202

Patton, E.G., Sullivan, P.P., Moeng, C.-H. (2004) The influence of idealized heterogeneity on wet and dry planetary boundary layers coupled to the land surface. *J. Atmospheric Sciences* 62: 2078-2097

Pleim, J.E. Xiu, A. (1995) Development and Testing of a Surface Flux and Planetary Boundary Layer Model for Application in Mesoscale Models. *J. Appl. Meteorology* 34: 16 – 32

Pleim, J.E. (2007a) A Combined Local and Nonlocal Closure Model for the Atmospheric Boundary Layer. Part I: Model Description and Testing. *J. Appl. Meteor. Climatol.* 46: 1383–1395

Pleim, J.E. (2007b) A Combined Local and Nonlocal Closure Model for the Atmospheric Boundary Layer. Part II: Application and Evaluation in a Mesoscale Meteorological Model. *J. Appl. Meteor. Climatol.* 46: 1396–1409

Raupach, M. R. (2000) Equilibrium evaporation and the convective boundary layer. *Bound.-Layer Meteor.* 96: 107–141

Salathé, E. P., R. Steed, C. F. Mass, Zahn, P.H. (2008) A high-resolution climate model for the U.S. Pacific Northwest: Mesoscale feedbacks and local responses to climate change. *J. Climate* 21: 5708–5726

Sellers, P. J., Y. Mintz, Y. C. Sud, Dalcher A. (1986) A simple biosphere model (SiB) for use within general circulation models. *J. Atmos. Sci.* 43: 505-531

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, M. Duda, X.-Y. Huang, W. Wang and Powers, J.G. (2008) A Description of the Advanced Research WRF Version 3. NCAR Technical Report, http://www.mmm.ucar.edu/wrf/users/docs/arw_v3.pdf

Smirnova, T. G., J. M. Brown, Benjamin, S.G. (1997) Performance of Different Soil Model Configurations in Simulating Ground Surface Temperature and Surface Fluxes. *Monthly Weather Review* 125, 1870-1884

Stauffer, D. R., Seaman, N.L. (1990) Use of four-dimensional data assimilation in a limited-area mesoscale model. Part I: Experiments with synoptic-scale data. *Mon. Wea. Rev.* 118: 1250-1277

Stauffer, D. R., Seaman, N.L. (1994) On Multi-Scale Four-Dimensional Data Assimilation. *J. Appl. Meteor.* 33: 416-434

Storm, B., Basu, S. (2010) The WRF model forecast-derived low-level wind shear climatology over the United States Great Plains. *Energies* 3: 258-276

Sun S. Xue, Y. (2001) Implementing a new snow scheme in Simplified Simple Biosphere Model (SSiB), *Advance in Atmospheric Sciences* 18: 335-354

van den Berg, G.P. (2008) Wind turbine power and sound in relation to atmospheric stability. *Wind Energy* 11: 151-69

Xiu, A., Pleim, J.E. (2001) Development of a Land Surface Model. Part I: Application in a Mesoscale Meteorological Model. *J. Appl. Meteor.* 40: 192-209

Xue, Y., P. J. Sellers, J. L. Kinter, Shukla J. (1991) A simplified biosphere model for global climate studies. *J. Climate* 4: 345-364

Xue, Y., F. J. Zeng, Schlosser, C.A. (1996) SSiB and its sensitivity to soil properties – a case study using HAPEX-Mobilhy data. *Global and Planetary Change* 13: 183 – 194

Wagner, R., Antoniou, I., Pedersen, S. M., Courtney, M. S., Jørgensen, H. E. (2009) The influence of the wind speed profile on wind turbine performance measurements *Wind Energy* 12: 348-62

Weckwerth, T., Parsons, D.B., Koch, S.E., Moore, J.A., LeMone, M.A., Demoz, B.B., Flamant, C., Geerts, B., Wang, J., Feltz, W.F. (2004) An overview of the International H₂O Project (IHOP_2002) and some preliminary highlights. *Bull. Amer. Meteor. Soc.* 85: 253-277

Wharton S., Lundquist, J.K. (2012) Assessing atmospheric stability and its impact on rotor disk wind characteristics at an onshore wind farm. *Wind Energy* 15: 525-546

Wharton S., Lundquist, J.K. (2012) Atmospheric stability affects wind turbine power collection. *Environmental Research Letters* 7(1): DOI: 10.1088/1748-9326/7/1/014005

Wiser, R., Bolinger, M., Barbose, G., Darghouth, N., Hoen, B., Mills, A., Weaver, S., Porter, K., Buckley, M., Fink, S., Oteri, F., Tegen, S. (2013) 2012 Wind Technologies Market Report, DOE/GO-102013-3948, August 2013, 92 p.

